Journal of Medical Virology 56:44–51 (1998)

Isolation of a GB Virus-Related Genome From a Chimpanzee

Larry G. Birkenmeyer,* Suresh M. Desai, A. Scott Muerhoff, Thomas P. Leary, John N. Simons, Carla C. Montes, and Isa K. Mushahwar

Virus Discovery Group, Experimental Biology Research, Abbott Laboratories, North Chicago, Illinois

Recently, two new flaviviruses, GB virus A (GBV-A) and GB virus B (GBV-B), were identified in the plasma of a tamarin infected with the hepatitis GB agent. A third virus, GB virus C (GBV-C), was subsequently identified in humans. In the current study, representational difference analysis (RDA) was used to search for a new virus in the serum of a chimpanzee that developed acute resolving hepatitis following inoculation with a pool of chimpanzee plasma. The plasma pool originated from serial passages of a human sample containing virus-like particles. Numerous cDNA clones were obtained that exhibited 62–80% identity with GBV-C. With the exception of the extreme 5′ and 3′ ends, the complete viral genome was sequenced, revealing a single large open reading frame encoding a 2833 amino acid polyprotein that contains two envelope proteins, two proteases, a helicase, and an RNA-dependent RNA polymerase. Phylogenetic analysis of the new virus indicates that it is closely related to GBV-C, yet still sufficiently divergent as to be placed in a separate group, tentatively labeled GB virus $C_{troglodytes}$ (GBV-C_{tro}). Numerous human samples were screened by reverse transcriptase-polymerase chain reaction (RT-PCR), but GBV-C_{tro} sequence was not detected. However, a second chimpanzee inoculated with the same plasma pool was shown to develop a GBV-C_{tro} infection. Although isolated from an Old World primate with hepatitis, the primary host of GBV-C_{tro} and any association with disease remains to be determined. ***J. Med. Virol. 56:44–51, 1998.***

KEY WORDS: GB virus; Chimpanzee; HCV

INTRODUCTION

In the search for new human hepatitis viruses, three novel flaviviruses, GBV-A, -B, and -C, related to HCV were described recently [Simons et al., 1996a]. The subtractive polymerase chain reaction (PCR) methodology known as representational difference analysis (RDA) was used to clone sequences of GBV-A and GBV-B from serum of a tamarin infected with the GB hepatitis agent [Simons et al., 1995b], while GBV-C was identified by reverse transcriptase-polymerase chain reaction (RT-PCR) in humans using GBV-A, GBV-B, and HCV helicase gene consensus primers [Simons et al., 1995a]. Although GBV-B appears to cause hepatitis in New World monkeys [Schlauder et al, 1995; Bukh et al., 1997], it has not been detected in humans. Conversely, infection with GBV-C accounts only for a few cases of human hepatitis. The existence of additional viral agents is suggested by several clinical studies of patients with elevated serum aminotranferase (ALT) levels [Marcellin et al., 1993], community-acquired acute resolving and chronic hepatitis [Alter et al., 1992; Tassopoulos et al., 1992; Buti et al., 1994], and fulminant hepatitis [Fagan et al., 1992; Kuwada et al., 1994; Sallie et al., 1994].

Potential starting points in the search for such a virus include reports of hepatitis caused by blood products derived from large pools of human plasma. A fibrinogen fraction, purified from such a pool, was reported to cause non-A/non-B hepatitis both in human recipients and in chimpanzees [Yoshizawa et al., 1980]. Convalescent serum from one of the infected chimpanzees was used to identify virus-like particles in the fibrinogen fraction and in a number of healthy human blood donors by immune electron microscopy. Acute hepatitis was induced in two chimpanzees inoculated with serum from one of these donors, indicating the presence of a transmissible agent. Based on the successful cloning of GBV-A and GBV-B sequences [Simons et al., 1995b], RDA was used to search for other viruses in infectious serum derived from serial chimpanzee passages of this human donor sample. We describe now a new member of the GB virus family present in serum from one of these chimpanzees, along with sequence analysis of the nearly complete genome.

*Correspondence to: Larry G. Birkenmeyer, Virus Discovery Group, Experimental Biology Research, Abbott Diagnostics Division, Abbott Laboratories, North Chicago, IL, 60064.

Accepted 7 April 1998

MATERIALS AND METHODS

Serum Samples

Previously, serum from a human donor containing virus-like particles was inoculated into a chimpanzee resulting in acute hepatitis [Yoshizawa et al., 1980]. Serum from this chimpanzee was passed serially through five more chimpanzees. A plasma pool from these five chimpanzees was inoculated into two additional chimpanzees, CH19, which developed acute resolving hepatitis, and CH413, which developed chronic hepatitis (data not shown).

Chimpanzee (*Pan troglodytes*) serum samples used in this study were obtained from CH19, CH413, six naive chimpanzees, and nine chimpanzees previously inoculated with potentially infectious material derived from human sources. Macaque (*Macaca fasicularis*) serum samples (18 in total) were obtained from captive animals. Human serum samples (276 in total) included normal and commercial blood donors, intravenous drug users, and non-A–E hepatitis patients. These individuals were from North America, Europe, and Asia.

The work with nonhuman primates was approved by the institution's ethical committee in accordance with established guidelines.

RDA

RDA was carried out essentially as described previously [Lisitsyn et al., 1993] with some modifications. Driver amplicon was derived from a serum pool (100 μl) obtained by combining equal volumes of serum from six naive chimpanzees. The pool was extracted (USB Total Nucleic Acid Extraction Kit) and cDNA was prepared prior to generation of the amplicon. Tester amplicon was prepared from two pools of CH19 serum-derived total nucleic acid, 0 to 30 days postinfection (pool 2) and 36 to 83 days postinfection (pool 3), as described for the driver, except that 100 μl of each tester pool was processed individually and then combined to generate a single tester amplicon. Three rounds of subtractive hybridization were performed and products from the third round were cloned into pT7 Blue (Novagen) for further analysis.

Genome Extension

The viral genomic sequence was determined by RT-PCR experiments utilizing cDNAs prepared from total nucleic acid extracted from CH19 pool 2 or pool 3 serum (USB Total Nucleic Acid Extraction Kit). Different combinations of RDA clone-specific primers were used to extend and connect the RDA-derived sequences where possible [Muerhoff et al., 1995]. In other experiments, PCR products were obtained by the "touchdown" PCR method [Roux, 1994], which used combinations of GBV-C-specific primers and RDA clone-specific primers for GBV-C_{tro}, particularly within regions of conservation as compared to GBV-C. Additional sequences were obtained using a PCR "walking" technique described by Sorensen et al. [1993].

DNA Sequencing and Sequence Analysis

PCR products were separated by electrophoresis through a 2% agarose gel and then excised and purified using the QIAEX Gel Extraction Kit (Qiagen, Chatsworth, CA). The purified fragments were either sequenced directly on an ABI Model 373 DNA Sequencer using the ABI Sequencing Ready Reaction Kit (Perkin-Elmer, Norwalk, CT) and gene-specific primers, or cloned into the pGEM-T Easy vector (Promega, Madison, WI) for sequence analysis. In cases where insufficient PCR product was obtained for direct sequencing, the amplicons were cloned and two or three clones from each were sequenced and consensus sequences were then generated. Individual sequences were compiled into a consensus sequence using Sequencher version 3.0 software (Gene Codes). Analysis of GBV-C_{tro} nucleotide and deduced polyprotein amino acid sequences was undertaken using the programs of the Wisconsin Package software (Wisconsin Sequence Analysis Package, Genetics Computer Group, Madison, WI). Alignments were made using the program PILEUP or the program ESEE, version 1.09c (Eric Cabot) based on GenBank accession numbers U22303 (GBV-A), U94421 (GBV-A_{lab}), AF023424 (GBV-A_{mx}), AF023425 (GBV-A_{tri}), U22304 (GBV-B), U36380 (GBV-C), AF070476 (GBV-C_{tro}), M62321 (HCV-1, genotype 1a), X61596 (HCV-JK1, genotype 1b), D00944 (HCV-J6, genotype 2a), D10988 (HCV-J8, genotype 2b), D28917 (HCV-3A, genotype 3a), and D26556 (HCV-3B, genotype 3b).

Evolutionary distances between sequences were determined using the DNADIST or PROTDIST program of the PHYLIP package, version 3.5c [Felsenstein, 1993]. The computed distances were used for the construction of phylogenetic trees using the neighbor-joining method of the program NEIGHBOR. Robustness of the trees was assessed by bootstrap resampling of the multiple sequence alignments using the programs SEQBOOT, DNADIST, and NEIGHBOR. The consensus tree was calculated by CONSENSE. Bootstrap values of less than 70% are regarded as not providing evidence for the phylogenetic grouping. The final graphical output was created with the program TREEVIEW [Page, 1996].

RT-PCR Screening Assay

A PCR assay was developed to detect GBV-C_{tro} RNA in the serum of humans or nonhuman primates. RT-PCR was carried out using the primers ntrC-3F and ntrC-4R essentially as described [Muerhoff et al., 1996a]. PCR products were separated by electrophoresis through 2% agarose gels followed by capillary transfer of nucleic acids to nylon membranes (Hybond-N+, Amersham, Arlington Heights, IL). The resulting blots were hybridized using a ^{32}P-end-labeled GBV-C_{tro}-specific oligonucleotide probe (5′-GTGCCGCCTGCGTTGTCGGGGTCTGGTATT-3′) under stringent conditions [Sambrook et al., 1989].

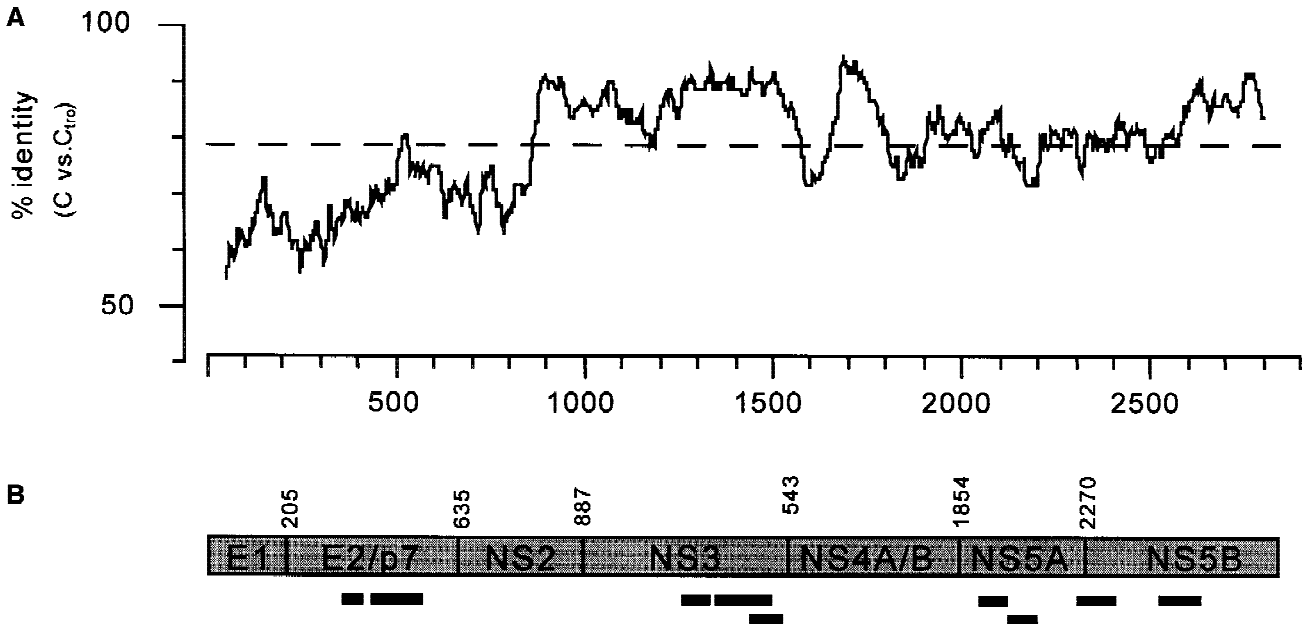

Fig. 1. **A:** Amino acid sequence identity of the GBV-C_{tro} polyprotein as compared to that of GBV-C. Following amino acid sequence alignment of the GBV-C_{tro} and GBV-C large open reading frames, the percentage identity of a 100-amino acid window of comparison was plotted. The dashed line represents the mean identity across the predicted polyprotein, and the amino acid scale is shown across the bottom. **B:** Genomic map of the GBV-C_{tro} polyprotein. The shaded box represents the GBV-C_{tro} large open reading frame, which begins at position 1 with the conserved methionine residue and concludes with the termination codon. Numbers above the polyprotein designate the putative first amino acid of the viral protein following proteolytic cleavage. The lines beneath the polyprotein represent the positions of the initial cDNA clones obtained by RDA. Alignments and plots are as described in Materials and Methods.

RESULTS AND DISCUSSION

Isolation of GBV-C_{tro}

Previously, antibodies to HCV were detected in the serum of CH19 and CH413, and RT-PCR southern blot analysis confirmed the presence of HCV sequences in CH19 serum (data not shown). Since it was unclear if the virus-like particles reported in the original human donor serum were associated with HCV, and since RDA was already shown to be capable of isolating sequences of more than one virus from a single sample [Simons et al., 1995b], it was decided to use RDA to search for additional viruses.

A preinoculation serum sample was not available for CH19, so the driver used for RDA subtraction was prepared from a serum pool from naive chimpanzees. After three rounds of enrichment, the RDA products were cloned and sequenced. Database searches revealed that 54 of the 65 inserts analyzed contained sequences related to GBV-C. The insert sequences represent nearly 25% of the GBV-C genome, including parts of the coding regions for E2, NS3, NS4, NS5A, and NS5B (Fig. 1B).

Even though HCV was present in the tester, none of the analyzed clones contained HCV sequences. One possible explanation is that the GBV-C-related virus titer appears to be higher than that of HCV in the CH19 serum, resulting in competitive loss of the HCV sequences during enrichment. Although the relative titers of the HCV and GBV-C-related virus are not known, RT-PCR-specific detection of HCV suggests its titer is low (data not shown), while the high representation in RDA products of the GBV-C-related sequences after only three rounds of enrichment suggests its titer is high. Since both CH19 and CH413 were coinfected with HCV, the potential for GBV-C_{tro} infection to cause hepatitis remains unclear.

GBV-C_{tro} Polyprotein Analysis

After the isolation of specific regions of the GBV-C_{tro} genome by RDA, the remaining sequences were extended as described in Materials and Methods. Currently, GBV-C_{tro} is 9,250 nucleotides in length, and, based on nucleotide sequence comparisons, is 33 nucleotides short of the GBV-C prototype genome on the 5′ end and 55 nucleotides short on the 3′ end. The sequence is 70.6% identical to the GBV-C prototype virus and represents the most divergent isolate of GBV-C identified to date.

In order to identify the initiating Met residue of the large polyprotein, nucleotide sequences from the 5′ end of GBV-C_{tro} through the N-terminus of the E1 gene were compared to those of the GBV-C prototype (Fig. 2). This region was 66% identical at the nucleotide level and contains a potential initiator Met codon immediately upstream of the putative signal sequence for the first envelope protein (E1). This Met codon is in-frame with a long open reading frame (ORF) of GBV-C_{tro} and is homologous to the sites of translation initiation mapped in GBV-A and GBV-C [Simons et al., 1996b]. Interestingly, the ORF of GBV-C_{tro} extends upstream

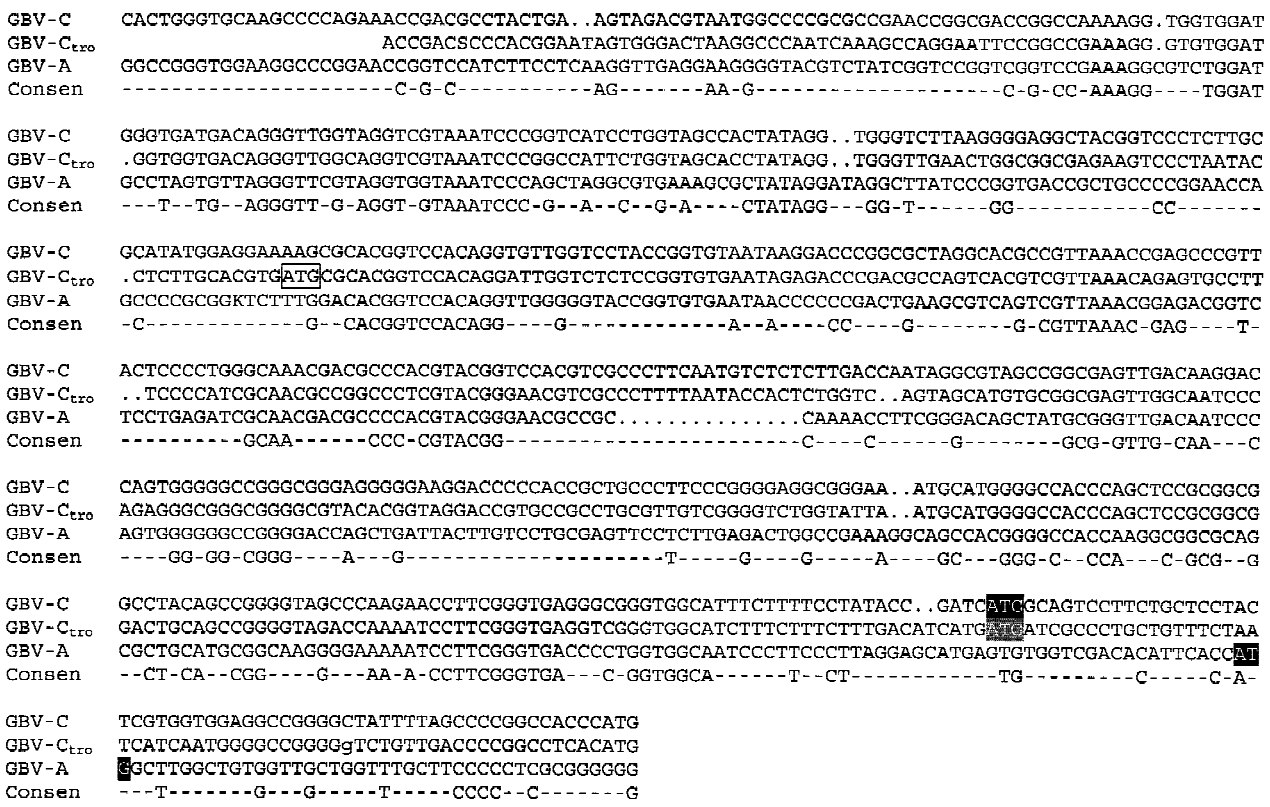

Fig. 2. Sequence alignment of the 5′ ends of GBV-C (nucleotides 13–597), GBV-C_{tro} (nucleotides 1–562), and GBV-A (nucleotides 59–637). Alignments were generated with the PILEUP program (default parameters) from GCG. The positions of the initiator Met codons for GBV-A and GBV-C [Simons et al., 1996b] are highlighted in black boxes. A homologous potential initiator codon for GBV-C_{tro} is highlighted in gray. Also shown is the position of the 5′ most ATG codon in-frame with the viral long ORF of GBV-C_{tro} (boxed).

from this codon potentially to encode a 109 residue peptide with an isoelectric point (pI) of 11.9. Although this is considerably shorter than the basic core proteins of GBV-B (156 amino acids) and HCV (191 amino acids), the pI of this peptide is in the range of viral core proteins (11.1 and 11.9 for GBV-B and HCV, respectively). Other GBV-C isolates that encode basic peptides upstream of E1 have been noted [Muerhoff et al, 1996b, 1997; Linnen et al., 1997; Takahashi et al., 1997]. However, in vitro translation reaction data suggest that these isolates initiate translation immediately upstream of E1 regardless of the presence of additional in-frame upstream sequences [Simons et al., 1996b]. Thus, GBV-C_{tro} appears to be similar to GBV-C, which does not encode a core protein and does not appear to contain a virion core [Simons et al, 1996b; Melvin et al, 1998]. Consistent with this possibility, the RNA secondary structure in this region that appears to be important for GBV-C's translation initiation is conserved in GBV-C_{tro} [Simons et al., 1996b; data not shown]. However, determining whether the ORF upstream of E1 is translated in GBV-C_{tro}-infected hosts will require additional study.

Utilizing the residue previously identified as the initiating Met in GBV-C [Simons et al., 1996b], the GBV-C_{tro} large ORF encodes a putative polyprotein of 2,833 amino acids. When the putative viral polyprotein cleavage products of GBV-C_{tro} are compared to those of GBV-C, extensive regions of identity can be found within several of the individual proteins (Fig. 1A). The greatest regions of identity occur within the nonstructural genes, NS3 through NS5B, while the lowest occurs in the envelope proteins. This variation in the structural proteins of GBV-C_{tro} may reflect a different host range relative to GBV-C. Within NS3, two regions of identity in excess of 90% occur at the termini of the protein corresponding to the protease and helicase domains, respectively [Leary et al., 1996b]. Closer examination of the helicase region within GBV-C_{tro} demonstrates that residues conserved in the supergroup II helicases of positive-strand RNA viruses [Koonin and Dolja, 1993] are maintained (Fig. 3A), indicative of a functional RNA helicase. A similar level of conservation is observed upon detailed examination of the GBV-C_{tro} NS5B region. Amino acids residues that define the supergroup II replicases of positive-strand RNA [Koonin and Dolja, 1993] are conserved within GBV-C_{tro}, including the Gly–Asp–Asp signature motif (residues 2576–2578) that is thought to be part of the catalytic domain of the replicase (Fig. 3B). Thus, in all probability, GBV-C_{tro} also encodes an RNA-dependent RNA polymerase similar to the other GB viruses and HCV.

An additional level of conservation that occurs between the GB viruses and HCV is in the two encoded

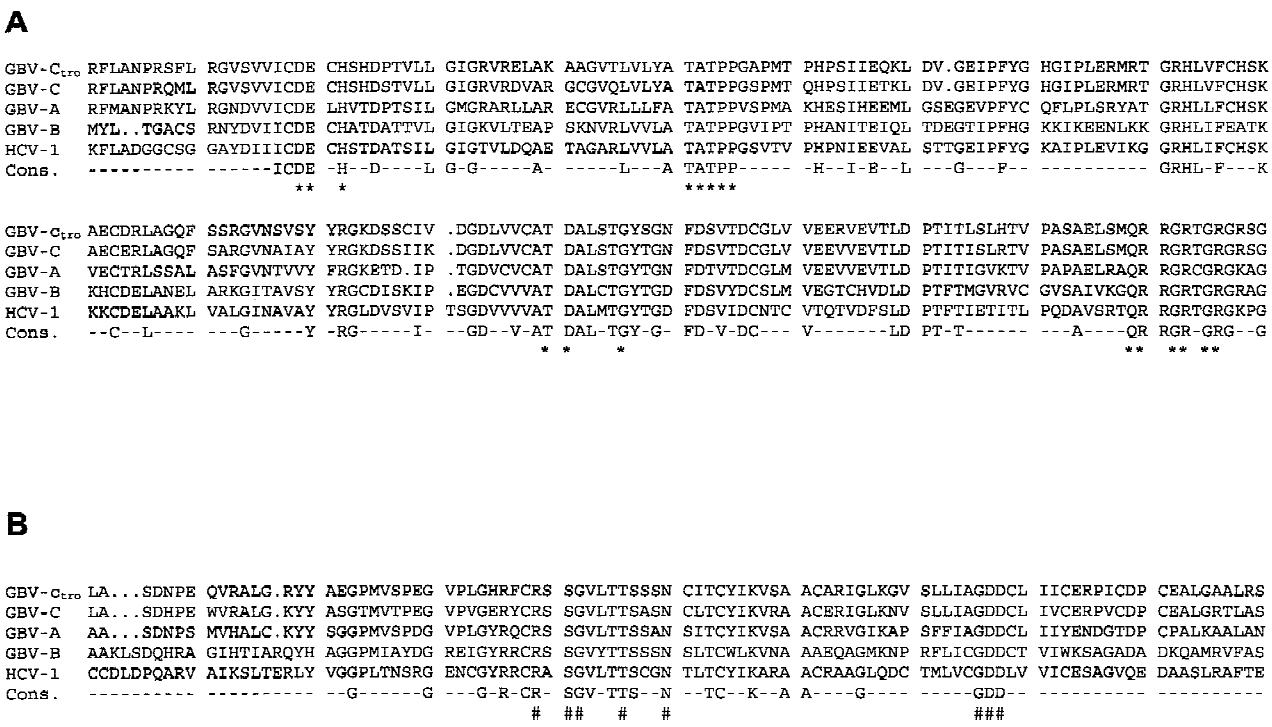

Fig. 3. Amino acid comparisons of helicase and replicase sequences. **A:** Alignment of the putative NS3 helicase domain of GBV-C_{tro} (residues 1159–1356), GBV-C (residues 1166–1363), GBV-A (residues 1234–1431), GBV-B (residues 1212–1408), and HCV-1 (residues 1298–1497). Residues conserved between the supergroup II RNA helicases are designated by asterisks. **B:** Alignment of the putative replicases around the GDD signature sequence of GBV-C_{tro} (residues 2505–2600), GBV-C (residues 2515–2610), GBV-A (residues 2626–2721), GBV-B (residues 2513–2612), and HCV-1 (residues 2662–2761). Residues conserved in the supergroup II replicases [Koonin and Dolja, 1993] are designated by number signs. Alignments were carried out as described in Materials and Methods.

proteases. At the N-terminus of the GBV-C_{tro} NS3 protein, within the protease region, the serine protease catalytic triad is conserved at positions 943 (His), 967 (Asp), and 1024 (Ser), as was the case in GBV-C and HCV [Leary et al., 1996b]. This suggests that GBV-C_{tro} processes the genomic polyprotein in a fashion similar to HCV, and in a manner that has been suggested for GBV-A, GBV-B [Muerhoff et al., 1995], and GBV-C [Leary et al., 1996b]. Supportive of this hypothesis is that the proteolytic cleavage sites that have been identified downstream of NS3 in GBV-B [Scarselli et al., 1997] and GBV-C [Belyaev et al., 1998] are also conserved in GBV-C_{tro}. The second encoded protease identified in GBV-C is a metal-dependent protease encoded within the NS2 region. Although the overall conservation between GBV-C_{tro} and GBV-C is lower within NS2, the His and Cys residues that are required for a functional protease [Grakoui et al., 1993] are conserved at positions 811 and 852, respectively. Additionally, the GBV-C proteolytic cleavage site between NS2 and NS3 that is hydrolyzed by this enzyme [Belyaev et al., 1998] is conserved in GBV-C_{tro} (with the scissile bond occurring between residues 887–888).

GBV-C_{tro} Host Range

To define more clearly the primary host, an RT-PCR oligonucleotide hybridization assay specific for GBV-C_{tro} was carried out on chimpanzee, macaque, and human samples. Samples from 16 captive chimpanzees were screened. Two serum samples were tested for each of nine chimpanzees, one prior to inoculation, and a second 3 to 12 months after inoculation with various human-derived hepatitis samples. A single serum sample was tested for each of the remaining seven chimpanzees for a total of 25 samples. Only the postinoculation serum from CH413 was positive by hybridization for the GBV-C_{tro} sequence. This animal had been inoculated with the same plasma pool as CH19. Sequencing of NS3 and NS5 PCR products from CH413 revealed nearly 100% identity with the GBV-C_{tro} sequence (data not shown). Using the same assay conditions, none of the 276 human or 18 macaque samples tested were positive for the presence of the GBV-C_{tro} virus.

Two lines of evidence suggest that GBV-C_{tro} may in fact be a chimpanzee virus. First, although most closely related to human GBV-C, the polyprotein sequence divergence between human GBV-C and GBV-C_{tro} is nearly four-fold that of all full length human GBV-C isolates. A similar pattern is observed for GBV-A, where divergence between isolates from within a species is significantly less than that between species. Second, the GBV-C_{tro} virus was not detected in an extensive screen of human samples, many of which were from patients with cryptogenic hepatitis or individuals at risk of contracting blood-borne viruses. Therefore, if

TABLE I. Percent Identity Between Viral Polyprotein Sequences[a]

	GBV-C	GBV-A_{lab}	GBV-A	GBV-A_{mx}	GBV-A_{tri}	HCV-J6	HCV-J8	HCV-1	HCV-JK1	HCV-3A	HCV-3B	GBV-B
GBV-C_{tro}	83.6*	54.4	54.2	54.9	54.6	31.7	32.4	31.8	31.7	32.0	32.2	29.9
GBV-C		54.3	54.1	54.4	54.4	32.2	32.4	32.7	32.4	32.5	32.4	29.6
GBV-A_{lab}			79.6*	78.1*	66.1*	32.4	32.8	33.3	32.7	32.6	32.3	29.8
GBV-A				78.1*	66.8*	32.2	32.1	32.0	31.9	32.0	31.6	29.2
GBV-A_{mx}					66.7*	32.3	32.7	32.5	32.4	32.7	32.8	29.5
GBV-A_{tri}						33.0	32.9	33.9	33.3	33.0	32.4	31.1
HCV-J6							88.4*	78.1*	77.1*	76.1*	74.4*	41.4
HCV-J8								78.0*	76.8*	75.9*	74.5*	41.8
HCV-1									88.5*	80.1*	78.8*	40.7
HCV-JK1										79.4*	78.1*	40.9
HCV-3A											88.2*	41.1
HCV-3B												39.8

[a]Amino acid sequence alignment included the entire polyprotein sequence of GBV-A and its variants, GBV-C, and GBV-C_{tro} (E1 through NS5B). HCV and GBV-B core sequences were not included since GBV-A, GBV-C, and GBV-C_{tro} do not contain homologous sequences. Asterisks indicate pairwise distances between sequences within the GBV-C, GBV-A, or HCV virus groups.

the diversity between GBV-C and GBV-C_{tro} reflects different human subtypes, similar to that observed between human HCV isolates, then GBV-C_{tro} must be quite rare in the human population. A source other than chimpanzee or human is unlikely since the only nonchimpanzee material traceable to the inoculum given to CH19 and CH413 is the original human donor serum. Failure of the RT-PCR assay to detect GBV-C_{tro} in human samples due to sequence variability cannot be ruled out. Additional GBV-C_{tro} isolates are needed to address this possibility and to establish firmly the primary host.

GBV-C Phylogenetic Analysis

As demonstrated by the pairwise distances shown in Table I, GBV-C_{tro} exhibited the highest percent identity with GBV-C (83.6%), followed by the GBV-A variants (54%), HCV (32%), and GBV-B (30%). Thus, GBV-C_{tro} is clearly most closely related to GBV-C. To determine more accurately the evolutionary relationship between GBV-C_{tro}, the other GB viruses, and HCV, we carried out phylogenetic analysis on their aligned 5′-untranslated regions (5′-UTR) and polyprotein sequences. As shown in Figure 4B, the phylogenetic groupings obtained from analysis of the putative polyprotein sequences demonstrate that the GBV-C/C_{tro}, the GBV-A variants, and the HCV genotypes form unique groups, each on one major branch of the tree. This grouping is confirmed by the very strong bootstrap support (100%). GBV-B forms a monophyletic group, as has been previously demonstrated [Muerhoff et al., 1995; Leary et al., 1996b]; however, polyprotein sequence comparisons indicate slightly greater identity with HCV than with GB viruses A or C. In addition, GBV-B and HCV share significant 5′-UTR sequence similarity and secondary structure [Simons et al., 1996b] and they both encode core proteins, whereas GBV-A and -C do not. Finally, the GB virus A and C groupings demonstrated by phylogenetic analysis of the polyproteins are identical to that obtained by 5′-UTR sequence analysis (Fig. 4A).

It is of interest to note the considerable genetic variability that exist within these virus groups. For example, the mean intragroup distances calculated from the polyprotein sequence alignments are very similar, i.e., approximately 73% for the GBV-A group, 84% for GBV-C/C_{tro}, and 80% for the HCV group. This similarity exists despite the fact that the GBV-A group consists of species-specific variants, whereas HCV variants are all human isolates. The significance of a similar level of diversity between GBV-C and GBV-C_{tro}, therefore, remains to be determined. Pairwise comparisons of 19 full-length polyprotein sequences of HCV representing the six major genotypes [Shukla et al., 1995] demonstrated a maximum divergence of approximately 30%. In contrast, pairwise comparisons of 16 full-length GBV-C polyprotein sequences demonstrated a maximum divergence of 4% (data not shown). GB viruses A and C, therefore, appear be to more evolutionarily constrained within their respective hosts than is HCV. Additionally, evolutionary distances at synonymous and nonsynonymous sites calculated within coding region sequences of HCV type 1 or subtype 1b demonstrated significantly greater variability than did GBV-C isolates [Muerhoff et al., 1997]. While multiple, genome-length sequences of GBV-A isolates from a single species are not available for comparison; analysis of 5′-UTR sequences suggests a degree of sequence conservation among species-specific GBV-A isolates analogous to that of GBV-C isolates [Leary et al., 1996a].

From these analyses it can be inferred that fundamental differences exist between these viruses with respect to their biology and the kinds of evolutionary pressures that are exerted on them. Recent studies have demonstrated that GBV-C antibodies directed against the second envelope protein, E2, appear to be neutralizing since anti-E2-positive individuals are rarely GBV-C RNA-positive [Pilot-Matias et al., 1996; Gutierrez et al., 1997; Tacke et al., 1997a; Ross et al., 1998]. This contrasts with HCV, where virus may persist in the presence of antibodies directed against both E2 and nonstructural proteins. Furthermore, HCV sequences obtained from persistently infected individuals exhibit a region of amino acid sequence hypervariablity within the E2 amino terminus [Weiner et al.,

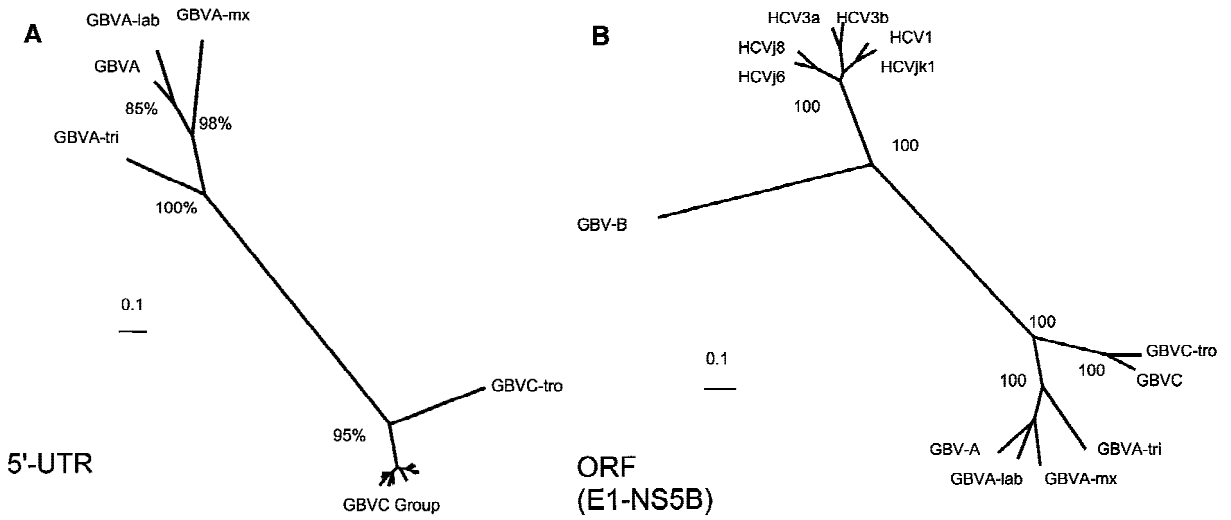

Fig. 4. Unrooted phylogenetic trees. Sequence alignment and phylogenetic analysis were as described in the Materials and Methods. **A:** 5′-UTR tree based on alignments of sequences from GBV-A variants, GBV-C variants, and GBV-C_{tro}. Sequences analyzed were from –499 to +89 of GBV-C_{tro}, where the adenine residue of presumed initiator ATG codon is designated as +1. The homologous region of the other viruses shown was included in the comparison. The bootstrap percentages based on 1,000 resamplings of the alignment data are shown at the appropriate nodes. GBV-C 5′-UTR sequences are from Muerhoff et al. [1996b]. **B:** Tree based on alignment of large open reading frame sequences from GBV-A variants, GBV-B, GBV-C variants, GBV-C_{tro}, and HCV genotypes. The sequences analyzed extend from the E1 protein amino terminus through the NS5B protein carboxy terminus. The core amino acid sequences of GBV-B and HCV were not included in the alignment since no homologous sequence exists in the large ORF of GBV-C, GBV-C_{tro}, or the GBV-A variants. The bootstrap percentages based on 100 resamplings of the alignment data are shown at the appropriate nodes.

1991]. This apparent fixation of mutations is believed to be due to host-immune selective pressures and the quasispecies nature of HCV [Booth et al., 1998]. No such E2 hypervariablity has been observed among GBV-C isolates [Erker et al., 1996; Tacke et al., 1997b], nor within an individual persistently infected with GBV-C over an eight-year period [Nakao et al., 1997]. Thus, while GBV-C and HCV can elicit a humoral immune responses in the infected host, they apparently use different mechanisms to establish long-lasting infections. It remains to be determined whether GB-A viruses or GBV-C_{tro} exhibit an anti-E2 response and whether the virus persists in the presence of anti-E2.

REFERENCES

Alter MJ, Margolis HS, Krawczynski K, Judson FN, Mares A, Alexander WJ, Hu PY, Miller JK, Gerber MA, Sampliner RE, Meeks EL, Beach MJ (1992): The natural history of community-acquired hepatitis C in the United States. New England Journal of Medicine 327:1899–1905.

Belyaev A, Chong S, Novikov A, Kongpachith A, Masiarz FR, Lim M, Kim JP (1998): Hepatitis G virus encodes protease activities which can effect processing of the virus putative nonstructural proteins. Journal of Virology 72:868–872.

Booth JCL, Kumar U, Webster D, Monjardino J, Thomas HC (1998): Comparison of the rate of sequence variation in the hypervariable region of E2/NS1 region of hepatitis C virus in normal and hypogammaglobulinemic patients. Hepatology 26:223–227.

Bukh J, Apgar CL, Purcell RH (1997): Natural history of GBV-A and GBV-B in primate animal models: Discovery of indigenous Flaviviridae-like viruses in several species of New World monkeys. In Rizzetto M, Purcell RH, Gerin JL, Verme G (eds). Viral Hepatitis and Liver Disease, Proceedings of IX Triennial International Symposium on Viral Hepatitis and Liver Disease, 21–25 April 1996. Rome, Italy: Edizioni Minerva Medica, pp 392–395.

Buti M, Jardi R, Rodriguez-Frias F, Quer J, Esteban R, Guardia J (1994): Etiology of acute sporadic hepatitis in Spain: The role of hepatitis C and E viruses. Journal of Hepatology 20:589–592.

Erker JC, Simons JN, Muerhoff AS, Leary TP, Chalmers ML, Desai SM, Mushahwar IK (1996): Molecular cloning and characterization of a GB virus C isolates from a patient with non-A–E hepatitis. Journal of General Virology 77:2713–2720.

Fagan EA, Ellis DS, Tovey GM, Lloyd G, Smith HM, Portmann B, Tan K-C, Zuckerman AJ, Williams R (1992): Toga virus-like particles in acute liver failure attributed to sporatic non-A, non-B hepatitis and recurrence after liver transplantation. Journal of Medical Virology 38:71–77.

Felsenstein J (1993): PHYLIP inference package, version 3.5c. Department of Genetics, University of Washington, Seattle.

Grakoui A, McCourt DW, Wychowski C, Feinstone SM, Rice CM (1993): A second hepatitis C virus-encoded proteinase. Proceedings of the National Academy of Sciences U.S.A. 90:10583–10587.

Gutierrez RA, Dawson GJ, Knigge MF, Melvin SL, Heynen CA, Kyrk CR, Young CE, Carrick RJ, Schlauder GG, Surowy TK, Dille BJ, Coleman PF, Thiele DL, Lentino JR, Pachucki C, Mushahwar IK (1997): Seroprevalence of GB virus C and persistence of RNA and antibody. Journal of Medical Virology 53:167–173.

Koonin EV, Dolja VV (1993): Evolution and taxonomy of positive-strand RNA viruses: Implications of comparative analysis of amino acid sequences. Critical Reviews in Biochemistry and Molecular Biology 28:375–430.

Kuwada SK, Patel VM, Hollinger B, Lin HJ, Yarbough PO, Wiesner RH, Kaese D, Rakela J (1994): Non-A, non-B fulminant hepatitis is also non-E and non-C. American Journal of Gastroenterology 89:57–61.

Leary TP, Desai SM, Yamaguchi J, Chalmers ML, Schlauder GG, Dawson GJ, Mushahwar IK (1996a): Species-specific variants of GB virus A in captive monkeys. Journal of Virology 70:9028–9030.

Leary TP, Muerhoff AS, Simons JN, Pilot-Matias TJ, Erker JC, Chalmers MC, Schlauder GG, Dawson GJ, Desai SM, Mushahwar IK (1996b): Sequence and genomic organization of GBV-C: A novel member of the *Flaviviridae* associated with human non-A–E hepatitis. Journal of Medical Virology 48:60–67.

Linnen JM, Fung K, Fry KE, Mizokami M, Ohba K, Wages JM, Zhang-Keck Z-Y, Song K, Kim JP (1997): Sequence variation and phylogenetic analysis of the 5′ terminus of hepatitis G virus. Journal of Viral Hepatitis 4:293–302.

Lisitsyn L, Lisitsyn N, Wigler M (1993): Cloning the differences between two complex genomes. Science 259:946–951.

Marcellin P, Martinot-Peignoux M, Gabriel F, Branger M, Degott C,

Elias A, Xu LZ, Larzul D, Erlinger S, Benhamou JP (1993): Chronic non-B, non-C hepatitis among blood donors assessed with HCV third-generation tests and polymerase chain reaction. Journal of Hepatology 19:167–170.

Melvin SL, Dawson GJ, Carrick RJ, Schlauder GG, Heynen CA, Mushahwar IK (1998): Biophysical characterization of GB virus C from human plasma. Journal of Virological Methods 71:147–157.

Muerhoff AS, Leary TP, Simons JN, Pilot-Matias TJ, Erker JC, Chalmers MC, Schlauder GG, Dawson GJ, Desai SM, Mushahwar IK (1995): Genomic organization of GBV-A and GBV-B: Two new members of the flaviviridae associated with GB-agent hepatitis. Journal of Virology 69:5621–5630.

Muerhoff AS, Simons JN, Erker JC, Desai SM, Mushahwar IK (1996a): Conserved nucleotide sequences within the GB Virus C 5′ untranslated region: Design of PCR primers for detection of viral RNA. Journal of Virological Methods 62:55–62.

Muerhoff AS, Simons JN, Leary TP, Erker JC, Chalmers ML, Pilot-Matias TJ, Dawson GJ, Desai SM, Mushahwar IK (1996b): Sequence heterogeneity within the 5′ terminal region of the hepatitis GB virus C genome and evidence for genotypes. Journal of Hepatology 25:379–384.

Muerhoff AS, Smith DB, Leary TP, Erker JC, Desai SM, Mushahwar IK (1997): Identification of GB virus C variants by phylogenetic analysis of 5′ untranslated and coding region sequences. Journal of Virology 71:6501–6508.

Nakao H, Okamoto H, Fukuda M, Tsuda F, Mitsui T, Masuko K, Iizuka H, Miyakawa Y, Mayumi M (1997): Mutation rate of GB virus C/Hepatitis G virus over the entire genome and in subgenomic regions. Virology 233:43–50.

Page RDM (1996): TREEVIEW: An application to display phylogenetic trees on personal computers. Computer Applications in the Biosciences 12:357.

Pilot-Matias TJ, Carrick RJ, Coleman PF, Leary TP, Surowy TK, Simons JN, Muerhoff AS, Buijk SL, Chalmers ML, Dawson GJ, Desai SM, Mushahwar IK (1996): Expression of the GB virus C E2 glycoprotein using the Semliki forest virus vector system and its utility as a serologic marker. Virology 225:282–292.

Ross RS, Viazov S, Schmitt U, Schmolke S, Tacke M, Ofenloch-Haehnle B, Holtmann M, Muller N, Villa GD, Yoshida CF, Oliveira JM, Szabo A, Paladi N, Kruppenbacher JP, Philipp T, Roggendorf M (1998): Distinct prevalence of antibodies to the E2 protein of GB virus C/hepatitis G virus in different parts of the world. Journal of Medical Virology 54:103–106.

Roux KH (1994): Using mismatched primer—template pairs in touchdown PCR. Bio/Techniques 16:812–814.

Sallie R, Silva AE, Purdy M, Smith H, McCaustland K, Tibbs C, Portmann B, Eddleston A, Bradley D, Williams R (1994): Hepatitis C and E in non-A non-B fulminant hepatic failure: A polymerase chain reaction and serological study. Journal of Hepatology 20:580–588.

Sambrook J, Fritsch EF, Maniatis T (1989). "Molecular Cloning: A Laboratory Manual." Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Scarselli E, Urbani A, Sbardellati A, Tomei L, Francesco RD, Traboni C (1997): GB virus B and hepatitis C virus NS3 serine proteases share substrate specificity. Journal of Virology 71:4985–4989.

Schlauder GG, Dawson GJ, Simons JN, Pilot-Matias TJ, Gutierrez RA, Heynen CA, Knigge MF, Kurpiewski GS, Leary TP, Muerhoff AS, Desai SM, Mushahwar IK (1995): Molecular and serologic analysis in the transmission of the GB hepatitis agents. Journal of Medical Virology 46:81–90.

Shukla DD, Hoyne PA, Ward CW (1995): Evaluation of complete genome sequences and sequences of individual gene products for the classification of hepatitis C viruses. Archives of Virology 140:1747–1761.

Simons JN, Leary TP, Dawson GJ, Pilot-Matias TJ, Muerhoff AS, Schlauder GG, Desai SM, Mushahwar IK (1995a): Isolation of novel virus-like sequences associated with human hepatitis. Nature Medicine 1:564–569.

Simons JN, Pilot-Matias TJ, Leary TP, Dawson GJ, Desai SM, Schlauder GG, Muerhoff AS, Erker JC, Buijk SL, Chalmers ML, Sant CLV, Mushahwar IK (1995b): Identification of two flavivirus-like genomes in the GB hepatitis agent. Proceedings of the National Academy of Sciences U.S.A. 92:3401–3405.

Simons JN, Desai SM, Mushahwar IK (1996a): The GB viruses: Isolation, characterization, diagnosis and epidemiology. Viral Hepatitis Reviews 2:229–246.

Simons JN, Desai SM, Schultz DE, Lemon SM, Mushahwar IK (1996b): Translation initiation in GB viruses A and C: Evidence for internal ribosome entry and implications on genome organization. Journal of Virology 70:6126–6135.

Sorenson AB, Duch M, Jorgensen P, Pedersen FS (1993): Amplification and sequence analysis of DNA flanking integrated proviruses by a simple two-step polymerase chain reaction method. Journal of Virology 67:7118–7124.

Tacke M, Kiyosawa K, Stark K, Schlueter V, Ofenloch-Haehnle B, Hess G, Engel A (1997a): Detection of antibodies to a putative hepatitis G virus envelope protein. Lancet 349:318–320.

Tacke M, Schmolke S, Schlueter V, Sauleda S, Esteban JI, Tanaka E, Kiyosawa K, Alter HJ, Schmitt U, Hess G, Ofenloch-Haehnle B, Engel AM (1997b): Humoral immune response to the E2 protein of hepatitis G virus is associated with long-term recovery from infection and reveals a high frequency of hepatitis G virus exposure among healthy blood donors. Hepatology 26:1626–1633.

Takahashi K, Hijikata M, Aoyama K, Hoshino H, Hino K, Mishiro S (1997): Characterization of GBV-C/HGV viral genome: Comparison among different isolates for a ~2kb sequences that covers entire E1 and most of 5′ UTR and E2. International Hepatology Communications 6:253–263.

Tassopoulos NC, Hatzakis A, Delladetsima I, Koutelou MG, Todoulos A, Miriagou V (1992): Role of hepatitis C virus in acute non-A, non-B hepatitis in Greece: A five-year prospective study. Gastroenterology 102:969–972.

Weiner AJ, Brauer MJ, Rosenblatt J, Richman KH, Tung J, Crawford K, Bonino F, Saracco G, Choo QL, Houghton M, Han JH (1991): Variable and hypervariable domains are found in the regions of HCV corresponding to the flavivirus envelope and NS1 proteins and the pestivirus envelope glycoproteins. Virology 180:842–848.

Yoshizawa H, Akahane Y, Itoh Y, Iwakiri S, Kitajima K, Morita M, Tanaka A, Nojiri T, Shimizu M, Miyakawa Y, Mayumi M (1980): Viruslike particles in a plasma fraction (fibrinogen) and in the circulation of apparently healthy blood donors capable of inducing non-A/non-B hepatitis in humans and chimpanzees. Gastroenterology 79:512–520.